```
          GenCore version 5.1.4_p5_4578
          Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   May  8, 2003, 14:51:24 ; Search time 6124 seconds
                                   (without alignments)
                                   13439.821 Million cell updates/sec

Title:           US-09-497-822C-18
Perfect score:   5082
Sequence:        1 gagctctgacaaattgag........acaagcaaacaaacaaaaaaaa 5082

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:        32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :  EST:*
            1:  em_estba:*
            2:  em_esthum:*
            3:  em_estin:*
            4:  em_estmu:*
            5:  em_estov:*
            6:  em_estpl:*
            7:  em_estro:*
            8:  em_htc:*
            9:  gb_est1:*
           10:  gb_est2:*
           11:  gb_htc:*
           12:  gb_est3:*
           13:  gb_est4:*
           14:  gb_est5:*
           15:  em_estfun:*
           16:  em_estom:*
           17:  gb_gss:*
           18:  em_gss_hum:*
           19:  em_gss_inv:*
           20:  em_gss_pln:*
           21:  em_gss_vrt:*
           22:  em_gss_fun:*
           23:  em_gss_mam:*
           24:  em_gss_mus:*
           25:  em_gss_other:*
           26:  em_gss_pro:*
           27:  em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	571.8	11.3	654	10	AW951855	AW951855 EST363925
2	529	10.4	532	9	AL704463	AL704463 DKFZp686A
3	519.4	10.2	634	10	BB612466	BB612466 BB612466
c 4	501.4	9.9	568	9	AA524966	AA524966 nh35c11.s
5	462.2	9.1	724	9	AI317423	AI317423 uj22e09.y
6	380.6	7.5	388	9	AA229063	AA229063 nc49g07.s
c 7	369	7.3	369	9	AA229062	AA229062 nc49g07.r
8	368.4	7.2	628	10	BB617970	BB617970 BB617970
9	361	7.1	413	14	T28396	T28396 EST41738 Hu
10	358.4	7.1	573	10	AW619116	AW619116 436 MARC
c 11	357.8	7.0	637	9	AI785019	AI785019 uj22e09.x
12	347.4	6.8	349	9	AA230070	AA230070 nc49b06.x
c 13	346	6.8	379	9	AA659567	AA659567 nt63a03.s
c 14	319	6.3	536	9	AI315085	AI315085 uj23g04.x
15	293.2	5.8	498	17	AZ017963	AZ017963 RPCI-23-2
c 16	285.4	5.6	292	9	AA229714	AA229714 nc49b06.r
17	247.6	4.9	395	17	AZ267276	AZ267276 RPCI-23-1
18	246.4	4.8	673	13	BI887100	BI887100 ZF637-1-0
c 19	238	4.7	247	9	AI659563	AI659563 tu12d12.x
20	234.2	4.6	396	10	AW619096	AW619096 424 MARC
21	226.8	4.5	557	13	BI979490	BI979490 ft88b09.y
22	221.4	4.4	664	9	AL709143	AL709143 DKFZp686L
23	216.8	4.3	667	13	BI183393	BI183393 UNL-P-FN-
24	214.6	4.2	413	9	AI326670	AI326670 mo60g04.y
c 25	203.4	4.0	378	9	AI893550	AI893550 mo60g04.x
26	196.2	3.9	322	9	AI317523	AI317523 uj23g04.y
27	193.8	3.8	1033	14	BQ049864	BQ049864 AGENCOURT
28	193.2	3.8	339	10	AW619117	AW619117 437 MARC
29	190.6	3.8	389	9	AA089319	AA089319 mo60g04.r
30	184.6	3.6	798	9	AU136010	AU136010 AU136010
31	178.4	3.5	582	14	BM714196	BM714196 UI-E-EJO-
32	178.4	3.5	647	14	BM714773	BM714773 UI-E-EJO-
33	174.6	3.4	613	13	BG938686	BG938686 cn27g02.x
34	173	3.4	643	12	BG345621	BG345621 dd89g10.y
35	171.8	3.4	621	12	BG886115	BG886115 dad73b08.
36	171.6	3.4	627	12	BG817008	BG817008 dad68h09.
37	167.6	3.3	441	10	BB850026	BB850026 BB850026
38	167.6	3.3	293	9	AA561201	AA561201 v128g06.r
39	164.6	3.2	546	13	BG997346	BG997346 MR4-HT105
40	163.6	3.2	667	10	BB666485	BB666485 BB666485
41	163.6	3.2	720	9	AU139958	AU139958 AU139958
42	162.4	3.2	831	12	BG614664	BG614664 602642157
43	159.4	3.1	521	10	AW910225	AW910225 ur79c10.y
c 44	158.4	3.1	464	17	AZ303301	AZ303301 UP_479-1J
45	157.6	3.1	459	9	AA252234	AA252234 zr63h08.r

ALIGNMENTS

```
RESULT 1
AW951855
LOCUS       AW951855                654 bp    mRNA    linear   EST 01-JUN-2000
DEFINITION  EST363925 MAGE resequences, MAGB Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW951855
VERSION     AW951855.1  GI:8141534
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 654)
AUTHORS     Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
            ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
TITLE       Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
JOURNAL     Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 32
            Seq primer: Reverse.
FEATURES             Location/Qualifiers
     source          1..654
                     /organism="Homo sapiens"
```

```
Fri May  9 14:20:41 2003                    us-09-497-822c-18.rst
```

```
Department Lehrach
Max-Planck-Institute for Molecular Genetics
Ihnestrasse 73, 14195 Berlin, Germany
Tel: +49-30-84131623
Fax: +49-30-84131128
Email: poustka@mpimg-berlin-dahlem.mpg.de
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No s1 sequence available.
This clone (DKFZp686A1931) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES             Location/Qualifiers
     source          1..532
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686A1931"
                     /clone_lib="686 (synonym: hlcc3)"
                     /tissue_type="human skeletal muscle"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     CDNA-collection"

BASE COUNT      139 a     131 c     142 g     117 t      3 others
ORIGIN
```

```
Query Match           10.4%;       Score 529;   DB 9;   Length 532;
Best Local Similarity 99.4%;       Pred. No. 1.5e-65;
Matches  529;  Conservative   0;   Mismatches   3;   Indels   0;   Gaps   0;

Qy  3634  GCTGAGGGAAACAGAAGTACCTGTGCGCCAGCAGAAATGATTGCACTATTGATAAATTC  3693
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  GCTGAGGGAAACAGAAGTACCTGTGCGCCAGCAGAAATGATTGCACTATTGATAAATTC    60

Qy  3694  CGAAGGAAAATTGTCCATCTTGCGTCTTCGGAAATGTTATGAAGCAGGGATGGACTCTG  3753
          | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  CGAAGGAAAATTGTCCATCTTGCGTCTTCGGAAATGTTATGAAGCAGGGATGGACTCTG   120

Qy  3754  GGAGCCCGGAAGCTGAAGAAACTTGGTAATCTGAAACTACAGGAGGAAGGAGAGGCTTCC  3813
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GGAGCCCGGAAGCTGAAGAAACTTGGTAATCTGAAACTACAGGAGGAAGGAGAGGCTTCC   180

Qy  3814  AGCACCACCAGCCCCACTGAGGAGACAACCCAGAAGCTGACAGTGTCACACATTGAAGGC  3873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  AGCACCACCAGCCCCACTGAGGAGACAACCCAGAAGCTGACAGTGTCACACATTGAAGGC   240

Qy  3874  TATGAATGTCAGCCCATCTTTCTGAATGTCCTGGAAGCCATTGAGCCAGTGTAGTGTGT  3933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  TATGAATGTCAGCCCATCTTTCTGAATGTCCTGGAAGCCATTGAGCCAGTGTAGTGTGT   300

Qy  3934  GCTGGACACGACACAACAACCAGCCCGACTCCTTGCTCCTTGCTCCTCTAGCCTCAATGAA  3993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  GCTGGACACGACACAACAACCAGCCCGACTCCTTGCTCCTTGCTCCTCTAGCCTCAATGAA   360

Qy  3994  CTGGGAGAGAGACAGCTTGTACACGTGGTCAAGTGGGCCAAGGCCTTGCCTGCCTCCGC  4053
          ||||||||||||||||||||||||||||||||||||  ||||| ||:|:||||:|||||
Db   361  CTGGGAGAGAGACAGCTTGTACACGTGGTCAAGTGGCCAAGGCCTTGCCTGCCTCCGC   420

Qy  4054  AACTTACACGTGGACGACCAGATGGCTGTCATTCAGTACTCCTGGATGGGCTCATGGTG  4113
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  AACTTACACGTGGACGACCAGATGGCTGTCATTCAGTACTCCTGGATGGGCTCATGGTG   480

Qy  4114  TTTGCCATGGGCTGGCGATCCTCCACCAATGTCAACTCCAGGATGCTCTACT  4165
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   481  TTTGCCATGGGCTGGCGATCCTCCACCAATGTCANNNCCAGGATGCTCTACT   532
```

```
RESULT 2
AL704463
LOCUS       AL704463          532 bp   mRNA    linear    EST  22-MAR-2002
DEFINITION  DKFZp686A1931_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686A1931 5', mRNA sequence.
ACCESSION   AL704463
VERSION     AL704463.1    GI:19687818
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 532)
AUTHORS     Poustka,A., Wellenreuther,R., Mewes,H.W., Weil,B. and Wiemann,S.).
TITLE       EST (Poustka,A., Wellenreuther,R., Mewes,H.W., Weil,B. and Wiemann
            ,S.)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Poustka A.J.
```

```
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGB"
                     /note="Vector: pBluescriptSKm"

BASE COUNT      166 a     183 c     133 g     172 t
ORIGIN

Query Match           11.3%;       Score 571.8; DB 10;   Length 654;
Best Local Similarity 95.3%;       Pred. No. 1.2e-71;
Matches  622;  Conservative   0;   Mismatches  27;   Indels   4;   Gaps   3;

Qy  4048  TTCCGCAACTTACACGTGACGACCAGATGGCTGTCATTCAGTACTCCTGGATGGGGCTC  4107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  TTCCGCAACTTACACGTGACGACCAGATGGCTGTCATTCAGTACTCCTGGATGGGGCTC    60

Qy  4108  ATGGTGTTTGCCATGGGCTGGCGATCCTCCACCAATGTCAACTCCAGGATGCTCTACTTC  4167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  ATGGTGTTTGCCATGGGCTGGCGATCCTCCACCAATGTCAACTCCAGGATGCTCTACTTC   120

Qy  4168  GCCCCTGATCTGGTTTTCAATGAGTACCGCATGCCACAAGTCCGGATGTACAGCCAGTGT  4227
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GCCCCTGATCTGGTTTTCAATGAGTACCGCATGCCACAAGTCCGGATGTACAGCCAGTGT   180

Qy  4228  GTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCTCCAAATCACCCCCAGGAATTC  4287
          |||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db   181  GTCCGAATGAGGCACCTCTCTCAAGAGTTTGGATGGCTCCAAATCACCCCCAGGAATTC   240

Qy  4288  CTGTGCATGAAAGCACTGCTACTCTTTCAGCATTATTCCAGTGGATGGGCTGAAAAATCAA  4347
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241  CTGTGCATGAAAGCACTGCTACTCTTTCAGCATTATTCCAGTGGATGGGCTGAAAAATCAA   300

Qy  4348  AAATTCTTTGATGAACTTCGAATGAACTACATCAAGGAACTCGATCGTATCATTCGCATGC  4407
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301  AAATTCTTTGATGAACTTCGAATGAACTACATCAAGGAACTCGATCGTATCATTCGCATGC   360

Qy  4408  AAAGAAAAATCCCACCTGCTCAAGACGCTTCTACCAGCTCCAAGCTCCTCGGAC  4467
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361  AAAGAAAAATCCCACCTGCTCAAGACGCTTCTACCAGCTCCAAGCTCCTCGGAC   420

Qy  4468  TCCGTGCAGCCTATTGCGAGAGAGCTCCATCAGTTCACTTTTGACCTGCTAATCAAGTCA  4527
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421  TCCGTGCAGCCTATTGCGAGAGAGCTCCATCAGTTCACTTTTGACCTGCTAATCAAGTCA   480

Qy  4528  CACATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGAGATCATCTCGTGTGCAAGTGCCC  4587
          ||||||||||||||||||| |||||||||||||||||||| ||||||||||||||||||||
Db   481  CACATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGAAATCATCTCGTGTGCAAGTGCCC   540

Qy  4588  AAGATCCTTTCTGG-AAAGTCAAGCCCATCTATTT-CCACACCCAGTGAAGCATTGGAA  4645
          ||||||||||| ||| ||||||||||||||||| ||  ||||||||||||||||| ||||
Db   541  AAGATCCTTTTGGGAAAAGTCAAGCCCCTTCTTTTCCCACACCCAGTGAAACCTTGGAA   600

Qy  4646  A--CCCTATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATGTCTTCTGCCTG  4696
          |  ||||||||||||||||||||| |  |||  |||
Db   601  AACCCCTTTTTCCCCACCCCCATTTTATGCCCCCCTTGAAAAAGTCTTTCTG   653
```

```
LOCUS        BB612466                634 bp    mRNA    linear   EST 31-AUG-2001
DEFINITION   BB612466 RIKEN full-length enriched, 0 day neonate skin Mus
             musculus cDNA clone 4632412C07 5', mRNA sequence.
ACCESSION    BB612466
VERSION      BB612466.1  GI:15394854
KEYWORDS     EST.
SOURCE       house mouse.
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 634)
  AUTHORS    Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
             , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
             ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
             Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
             ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
             Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
             Muramatsu,M. and Hayashizaki,Y.
  TITLE      RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL    Unpublished (2001)
COMMENT      Contact: Yoshihide Hayashizaki
             Laboratory for Genome Exploration Research Group, RIKEN Genomic
             Sciences Center(GSC), Yokohama Institute
             The Institute of Physical and Chemical Research (RIKEN)
             1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
             Tel: 81-45-503-9222
             Fax: 81-45-503-9216
             Email: genome-res@gsc.riken.go.jp.
             URL:http://genome.gsc.riken.go.jp/
             Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
             ,M., Konno,H., Okazaki,Y., Muramatsu.M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
             prepare full-length cDNA libraries for rapid discovery of new
             genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
             Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
             ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
             Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
             sequencing pipeline with 384 multicapillary sequencer. Genome Res.
             10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
             ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
             encyclopedia: real-time sequence clustering for construction of a
             nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa
             ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T.,
             Ishii,Y. and Hayashizaki,Y.
             Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
             Func. Genomics 2 pre, L72-L86 (2001
             Please visit our web site (http://genome.gsc.riken.go.jp) for
             further details.
FEATURES     Location/Qualifiers
     source  1..634
             /organism="Mus musculus"
             /strain="C57BL/6J"
             /db_xref="taxon:10090"
             /clone="4632412C07"
             /clone_lib="RIKEN full-length enriched, 0 day neonate
             skin"
             /sex="mixed"
             /tissue_type="skin"
             /dev_stage="0 day neonate"
             /lab_host="DH10B"
             /note="Site_1: SalI; Site_2: BamHI; cDNA library was
             prepared and sequenced in Mouse Genome Encyclopedia
             Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in
             RIKEN. Division of Experimental Animal Research in Riken
             contributed to prepare mouse tissues. 1st strand cDNA was
             primed with a primer [5'
```

```
GAGAGAGAGAGGATCCCAAGAGCTCTTTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 100.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCGAGTTAATTAATAATCCCCCCCCCCC
3']. cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified pBluescript KS(+) after bulk excision
from Lambda FLC I"

BASE COUNT       169 a    140 c    165 g    160 t
ORIGIN

Query Match             10.2%;    Score 519.4;   DB 10;   Length 634;
Best Local Similarity   88.8%;    Pred. No. 3.1e-64;
Matches 562;  Conservative  0;  Mismatches  71;  Indels  0;  Gaps  0;

Qy  3642 GAAACAGAAGTACCTGTGCGCCAGCAGAAATGATTGCACTATTGATAAATTCCGAAGGAA 3701
         ||||||||||| || ||||||| || ||| |||||| ||| | ||||||||||||||||
Db     1 GAAACAGAAGTATCTATGTGCCAGCAGAAACGATTGACCATTGATAAATTTCGAGGAA 60

Qy  3702 AAATTGTCCATCTTGCGTCTTCGGAAATGTTATGAAGCAGGATGACTCTGGGAGCCCG 3761
         ||||||||||| ||||||||||| |||||||| |||||||||||||||||||||| ||
Db    61 AAATTGCCCATCTTGCGTCGTCCGGAAATGTTATGAAGCAGGATGACTGCTGGGAGCTCG 120

Qy  3762 GAAGCTGAAGAAACTTGGTAATCTGAAACTACAGGAGGAAGGAGAGGCTTCCAGCACCAC 3821
         ||||||||||||||||| |||| ||||| ||||||||||||||||||||||  |||
Db   121 TAAGCTGAAGAAACTTGGAAATCTAAAACTACAGGAGGAAGGAGAGAAAACTCCAATGCTGG 180

Qy  3822 CAGCCCCACTGAGGAGACAACCCAGAAGCTGACAGTGTCACATTGAAGGCTATGAATG 3881
         |||||||||||| |||||||||||||| ||||||||| || || ||||||||||||||
Db   181 CAGCCCCACTGAGGACACCATCCCAGAAGATGACTGTATCACACATTGAAGGCTATGAATG 240

Qy  3882 TCAGCCATCTTTCTGAAATGTCCTGGAAGCCATTGAGCCAGGTGTAGNGTGTCTGGACA 3941
         ||||||||||| |||||||||||| |||||||||||||||| | |||||||||||||||
Db   241 TCAGCCTATCTTTCTTAACGTCCTGGAAGCCATTGAGCCAGGAGTGGTGTGTCCGGACA 300

Qy  3942 CGACAACAACCAGCCCGACTCCTTTGCAGCCTTGCTCTCTGAACGCCTCAATGAACTGGGAGA 4001
         |||||||||| || |||| |||| ||||||||||||| ||||||||||| ||||||||
Db   301 TGACAACAACCAACCAGATTCCTTTGCTCCTTGTTATCTAGCCTCAATGAGCTTGGAGA 360

Qy  4002 GAGACAGCTTGTACACGTGCAAGTGGTCAAGTGGCCAAGGCCTTCCGCAACTTACA 4061
         ||| |||||| |||| ||||||||||||||||||||||||||||||||||||||
Db   361 GAGGCAGCTTGTGCATGTGGTCAAGTGGTCAAGTGGCCAAGGCCTTCCGGCAACTTGCA 420

Qy  4062 CGTGGACGACCAGATGGCTGTCATTCAGTACTCCTGGATGGGCTCATGGTGTTGCCAT 4121
         |||| ||||||||||||||||| |||||||||||||||| |||||||||||||||||
Db   421 TGTTGGATGACCAGATGGGCGGTCATTCAGTATTCCTGGATGGGCGACTGATGGTAGTTGCCAT 480

Qy  4122 GGGCTGGGATCCTTCACCAATGTCACCAATGTCACAAGTCCCGGATGTACAGCCAGTGTGTCCGAATGAGGCA 4181
         |||| |||| |||| | |||||| |||| |||| || ||||| ||| || |||||||
Db   481 GGGTTGGCGGTCCTTCACTAATGTCAACTCCAGGATGTCCAGGATGTCTCACTTTGCACCTGACTTGGT 540

Qy  4182 TTTCAATGAGTACCGCCATGCACAAGTCCGGATGTACAGCCAGTGTGTCCGAATGAGGCA 4241
         |||||||||||||||| ||| ||||||||| ||| ||||||||||||||| ||||||||
Db   541 TTTCAATGAGTACCGCCATGCACAAGTCTTGGATGTACAACCAGTGTGTGAGCATGAGGCA 600

Qy  4242 CCTCTCTCAAGAGTTTGGATGGCTCCAAATCAC 4274
         || || |||||||||||||||||||| ||
Db   601 CCTGTCTCAAGAGTTTGGATGGCTCCAAATAAC 633
```

```
RESULT 4
AA524966/c
LOCUS        AA524966                568 bp    mRNA    linear  EST 05-AUG-1997
DEFINITION   nh35cl1.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:954356
             similar to gb:M23263 ANDROGEN RECEPTOR (HUMAN); mRNA sequence.
ACCESSION    AA524966
VERSION      AA524966.1  GI:2265894
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE   1 (bases 1 to 568)
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.
            , Michael Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 362   Std Error: 0.00
            Seq primer: ~40m13 fwd. ET from Amersham
            High quality sequence stop: 383.
FEATURES   Location/Qualifiers
     source 1..568
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:954356"
            /clone_lib="NCI_CGAP_Pr3"
            /sex="Male"
            /dev_stage="45 years old"
            /lab_host="DH10B"
            /note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st
            strand cDNA was primed with oligo(dT)17 on 50 ng of
            DNAse-treated, total cellular RNA obtained from 5,000-10
            ,000 microdissected cells histologically-determined to be
            fully malignant prostate cancer cells. Double-stranded
            cDNA was ligated to EcoRI adaptors, 5 cycles of PCR
            applied to the cDNA with an adaptor-specific primer, and
            the resulting PCR product subcloned into pAMP10 by the
            UDG-cloning method (Life Technologies). Average insert
            size is 600 bp. NOTE: Not directionally cloned. This
            library was constructed by David Krizman."
BASE COUNT   186 a   99 c   153 g   130 t
ORIGIN

Query Match   9.9%;   Score 501.4; DB 9;   Length 568;
Best Local Similarity   98.5%;   Pred. No. 1.2e-61;
Matches 527;   Conservative   0;   Mismatches   6;   Indels   2;   Gaps   2;

Qy   4529 ACATGGTGAGCGTGGACTTTCCGGAAATGATGCAGAGATCATCTCTGTGCAAGTG-CCC 4587
          |||||||||| |||||| |||||| ||||||| |||||||||||||||||||| |||
Db   550 ACATGGTGAGCGTGACTTTTCCGAAATGATGCCAGAGATCATCTCTGTGCAAGTGCCCC 491

Qy   4588 AAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTCCACACCCAGTGAAGCATTGAAAAC 4647
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db   490 AAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTCCACACCCAGTGAAGCA-TGGAAAC 432

Qy   4648 CCTATTCCCCACCCCAGCTCATGCCCCCTTTCAGATGTCTTCGCCTGTTATAACTCTG 4707
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   431 CCTATTCCCCACCCCAGCTCATGCCCCCTTTCAGATGTCTTCGCCTGTTATAACTCTG 372

Qy   4708 CACTACTCCTCTGCAGTGCCTTGGGAATTTCCTCTATTGATGTACAGTCTGTCATGAAC 4767
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   371 CACTACTCCTCTGCAGTGCCTTGGGAATTTCCTCTATTGATGTACAGTCTGTCATGAAC 312

Qy   4768 ATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTTTTTTCTCTTCTCCTTTCTTTTTCTTC 4827
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   311 ATGTTCCTGAATTCTATTTGCTGGGCTTTTTTTTTTTTCTCTTCTCCTTTCTTTTTCTTC 252

Qy   4828 TTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCTCCT 4887
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   251 TTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCTCCT 192

Qy   4888 ATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTGTGATGATCCTCATATGGCCCAG 4947
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   191 ATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAATCTGTGATGATCCTCATATGGCCCAG 132

Qy   4948 TGTCAAGTTGTGTGCTTGTTTACAGCACTACTCTGTGCAGCCACACAAACGTTTACTTATC 5007
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131 TGTCAAGTTGTGTGCTTGTTTACAGCACTACTCTGTGCAGCCACACAAACGTTTACTTATC 72

Qy   5008 TTATGCCACGGAAGTTTAGAGAGTCTAAGATTATATCTGGGAAATCAAACAAAA 5062
          |||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   71 TTATGCCACGGAAGTTTAGAGAGTCTAAGATTATATCTGAGGAAATCAAAACAAAA 17

RESULT 5
AI317423
LOCUS   AI317423   724 bp   mRNA   linear   EST 17-DEC-1998
DEFINITION uj22e09.y1 Sugano mouse kidney mkia Mus musculus cDNA clone
            IMAGE:1920712 5' similar to gb:X53779 Mouse mRNA for androgen
            receptor (MOUSE);, mRNA sequence.
ACCESSION   AI317423
VERSION   AI317423.1   GI:4032690
KEYWORDS   EST.
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 724)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE   The WashU-HHMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL.; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:977004
            Seq primer: custom primer used
            High quality sequence stop: 515.
FEATURES   Location/Qualifiers
     source 1..724
            /organism="Mus musculus"
            /strain="C57BL"
            /db_xref="taxon:10090"
            /clone="IMAGE:1920712"
            /clone_lib="Sugano mouse kidney mkia"
            /sex="female"
            /dev_stage="adult"
            /lab_host="DH10B"
            /note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
            (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
            was primed with an oligo(dT) primer
            [ATGTGGCCTTTTTTTTTTTTTTT]: double-stranded cDNA was
            ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
            and cloned into distinct DraIII sites of the pME18S-FL3
            vector (5' site CACCATGTG, 3' site CACCATGTG. XhoI should
            be used to isolate the cDNA insert. Size selection was
            performed to exclude fragments <1.5kb. Library
            constructed by Dr. Sumio Sugano (University of Tokyo
            Institute of Medical Science). Custom primers for
            sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end
            primer CGACCTGCAGCTGGAGCACA."
BASE COUNT   190 a   181 c   192 g   161 t
ORIGIN

Query Match   9.18;   Score 462.2; DB 9;   Length 724;
Best Local Similarity   82.88;   Pred. No. 3.6e-56;
Matches 564;   Conservative   0;   Mismatches 113;   Indels   4;   Gaps   3;

Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 366.

FEATURES   Location/Qualifiers
    source   1..388
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1011516"
                     /clone_lib="NCI_CGAP_Pr3"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated, total cellular RNA obtained from 5,000-10
                     ,000 microdissected cells histologically-determined to be
                     fully malignant prostate cancer cells. Double-stranded
                     cDNA was ligated to EcoRI adaptors, 5 cycles of PCR
                     applied to the cDNA with an adaptor-specific primer, and
                     the resulting PCR product subcloned into pAMP10 by the
                     UDG-cloning method (Life Technologies). Average insert
                     size is 600 bp. NOTE: Not directionally cloned. This
                     library was constructed by David Krizman."

BASE COUNT   77 a   113 c   72 g   126 t
ORIGIN

Query Match   7.5%;   Score 380.6; DB 9; Length 388;
Best Local Similarity 99.0%;   Pred. No. 1.7e-44;
Matches 383;   Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4442 TCTACCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGT 4501
          || ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   2 TCGACCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCTATTGCGAGAGAGCTGCATCAGT 61

Qy 4502 TCACTTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAATGATGG 4561
          ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   62 TCACTTTTGACCTGCTAATCAAGTCACACATGGTGAGCGTGGACTTTCCGGAAAATGATGG 121

Qy 4562 CAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATT 4621
          |||||||||||||||||||||||||||||||| ||||||||||||||||| ||||||||||
Db   122 CAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATT 181

Qy 4622 TCCACACCCAGTGAAGCATTGGAAACCCTATTCCCCACCCCAGCTCATGCCCCTTTCA 4681
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   182 TCCACACCCAGTGAAGCATTGGAAACCCTATTCCCCACCCCAGCTCATGCCCCTTTCA 241

Qy 4682 GATGTCTTCGCCTGTGTATAACTCTGCACTACTCCTGCACTACTCCTCTGCAGTGCCT 4741
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   242 GATGTCTTCGCCTGTGTATAACTCTGCACTACTCCTGCACTACTCCTCTGCAGTGCCT 301

Qy 4742 CTATTGATGTACAGTCGTCATGAACATGTTCCTGAATTCTATTTGCTGGCGTTTTTTT 4801
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302 CTATTGATGTACAGTCGTCATGAACATGTTCCTGAATTCTATTTGCTGGCGTTTTTTT 361

Qy 4802 TCTCTTTCTCTCCTCCTTTCTTTTTTCTTCT 4828
          ||||||||||||||||||||||||
Db   362 TCTCTTTCTCTCCTCCTTTCTTTTTTCTCT 388

RESULT 7
AA229062/c
LOCUS   AA229062   369 bp   mRNA   linear   EST 21-AUG-1997
DEFINITION nc49g07.r1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1011516
            similar to gb:M23263 ANDROGEN RECEPTOR (HUMAN);, mRNA sequence.
ACCESSION   AA229062
VERSION   AA229062.1 GI:1852046
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Qy 3257 GCGGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGCTGTAGCCCCTACG 3316
          || ||||||| || || | || | | | || | |||||||||||| |||||
Db   16 GCCTGGGCTGCGGCGGCAGC-GCAATGCCGCTATGGGGACTTGGGTAGTCAGCCCCTATG 74

Qy 3317 GCTACACTCGGCCCCCTCAGGGGCTGGCCGGCCAGGAAAGCGACTTCACCGCACCTGATG 3376
          |||||||||| ||||||||||||| ||||| |||||||| || ||| | | || | ||||
Db   75 GCTACACTCGGCCCCCTCAGGGGCTGACGAAGCCAGGAGTGACTACTCTGCCTCCGAAG 134

Qy 3377 TGTGTACCCTGGCGGCATGGTGAGCAGAGTGCCCTATCCCAGTTGTGTCAAAA 3436
          |||||| ||||| | |||| |||| |||||||||||||||||||||||||||||
Db   135 TGTGTATCCTGGTGGAGTTGAACAGATACCCTATCCCAGTTGTGTCAAAA 194

Qy 3437 GCGAAATGGGCCCTGGATGGATAGCTACTCCGGGACCTTACGGGGACATGCGTTTGGAGA 3496
          | |||||| | |||||||||| |||| ||||||| ||||||||||||||||||||||||
Db   195 GTGAAATGGGACCTTGGATGGAGAACTACTCCGGGACCCTACGGGGACACGCGTTTGGACA 254

Qy 3497 CTGCCAGGACCATTGTTTGCCCATTGACTATTACTTTCCACCCCAGAAGACCTGCCTGA 3556
          | ||||||||||||||| |||||||| || ||||| ||||||||||||||||||||||||
Db   255 GTACCAGGGACCATGTTTACCCATGGACTATTACTTTCCACCCCAGAAGACCTGCCTGA 314

Qy 3557 TCTGTGGAGATGAAGCTTCTGGTGTCACTATGGAGCTCTCACATGTGGAAGCTGCAAGG 3616
          |||||||||||||||||| ||||||||||||||||||||||||| || ||| ||||||||
Db   315 TCTGTGGAGATGAAGCTTCTGCTGTCACTACGGAGCTCTCACTTGTGCAGCTGCAAGG 374

Qy 3617 TCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCGCCCAGCAGAAATGATT 3676
          ||||||||||| |||| | ||||||||||||||||||||||||| || |||||||||||||
Db   375 TCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTATCTATGTGCCAGCAGAAACGATT 434

Qy 3677 GCACTATTGATAAATTCCGAAGGAAAAATTGTCCATCTTGTCGTCTTCGGAAATGTTATG 3736
          | ||||||||||||| | || || |||||||||| |||||||||| | |||||||||| |
Db   435 GTACCATTGATAAATTTCGGAGGAAAAATTGCCCATCTTGTCGTCTCCGGAAATGTTATG 494

Qy 3737 AAGCAGGAGGGATGACTCTGGGAGCCCGGAAGCTGAAGAAAACTGGTAATCTGAAACTACAGG 3796
          |||| ||||| || ||| | |||||| || ||||||||||||| |||||| | ||||||||||
Db   495 AAGCAAGGATGACTCTGGGAGCTCGTAAGCTGAAGAAACTTGGAAATCTAAACTACAGG 554

Qy 3797 AGGAAGGAGAGGCTTCAGCACCACCAGCCCACTGAGGAGACAAACCCAGAAGCTGACAG 3856
          ||||||||| ||| ||| ||||| ||||||||||||||||||| ||| ||||||||||
Db   555 AGGAAGGAGAAAACTCCAATGCTGCGGCACCCCACTGAGGACCCATCCAGGAAGATGACTG 614

Qy 3857 TGTCACACATTGAAGGC-TATGAAATGTCAGCCCATCTTT--CTGAAATGTCCTGGAAGCCA 3913
          | ||||| ||||||||| |||||||||| ||||||||||||| | ||||||||||||||||
Db   615 TATCACACATTGAAAGCTTTGAATGTCAGCCTATCTTTTCTTAACGTCCCTGGAAGCCA 674

Qy 3914 TTGAGCCAGGTGTAGTGTGTG 3934
          |||| ||||||||||| |||||
Db   675 TTTGAGCCCCAGGAATGGTGTG 695

RESULT 6
AA229063
LOCUS   AA229063   388 bp   mRNA   linear   EST 21-AUG-1997
DEFINITION nc49g07.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1011516
            similar to gb:M23263 ANDROGEN RECEPTOR (HUMAN);, mRNA sequence.
ACCESSION   AA229063
VERSION   AA229063.1 GI:1852047
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1 (bases 1 to 369)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.
            , Michael Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 341.
FEATURES   Location/Qualifiers
     source   1..369
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1011516"
                     /clone_lib="NCI_CGAP_Pr3"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNase-treated, total cellular RNA obtained from 5,000-10
                     ,000 microdissected cells histologically-determined to be
                     fully malignant prostate cancer cells. Double-stranded
                     cDNA was ligated to EcoRI adaptors, 5 cycles of PCR
                     applied to the cDNA with an adaptor-specific primer, and
                     the resulting PCR product subcloned into pAMP10 by the
                     UDG-cloning method (Life Technologies). Average insert
                     size is 600 bp. NOTE: Not directionally cloned. This
                     library was constructed by David Krizman."
BASE COUNT   142 a   67 c   93 g   67 t
ORIGIN

Query Match   7.3%; Score 369; DB 9; Length 369;
Best Local Similarity   100.0%; Pred. No. 7.9e-43;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   4680 CAGATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTC 4739
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   369 CAGATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCTGCAGTGCCTTGGGGAATTTC 310

Qy   4740 CTCTATTGATGTACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGCTTTTTTT 4799
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   309 CTCTATTGATGTACAGTCTGTCATGAACATGTTCCTGAATTCTATTTGCTGGCTTTTTTT 250

Qy   4800 TTTCTCTTTCTCTCCTTTCTTTTTTTCTTTTTCCTCCCTATCTAACCCTCCCATGGCACCT 4859
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   249 TTTCTCTTTCTCTCCTTTCTTTTTTTCTTTTTCCTCCCTATCTAACCCTCCCATGGCACCT 190

Qy   4860 TCAGACTTTGCTTCCCATTGTGGCCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTA 4919
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   189 TCAGACTTTGCTTCCCATTGTGGCCTCCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTA 130

Qy   4920 AATCTGTGATGATCCTCATATGGCCCAGTGTCAAGTTGCTTGTTGCTTGTTACAGCACTACTCT 4979
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   129 AATCTGTGATGATCCTCATATGGCCCAGTGTCAAGTTGCTTGTTGCTTGTTACAGCACTACTCT 70

Qy   4980 GTGCCAGCCACACAAACGTTTACTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATT 5039
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   69 GTGCCAGCCACACAAACGTTTACTTACTTATCTTATGCCACGGGAAGTTTAGAGAGCTAAGATT 10

Qy   5040 ATCTGGGGA 5048
          |||||||||
Db   9 ATCTGGGGA 1

RESULT 8
BB617970
LOCUS   BB617970   628 bp   mRNA   linear: EST 26-OCT-2001
DEFINITION BB617970 RIKEN full-length enriched, adult male pituitary gland Mus
            musculus cDNA clone 5330428G13 5'', mRNA sequence.
ACCESSION   BB617970
VERSION   BB617970.1 GI:16457679
KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 628)
AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
TITLE   RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL   Unpublished (2001)
COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
            wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
            RIKEN integrated sequence analysis (RISA) system-384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
            Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
            Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
            Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
            Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
            Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
            e mouse tissues.
FEATURES   Location/Qualifiers
     source   1..628
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="5330428G13"
                     /clone_lib="RIKEN full-length enriched, adult male
                     pituitary gland"
                     /sex="male"
                     /tissue_type="pituitary gland"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: Sall; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken

```
ACCESSION    T28396
VERSION      T28396.1    GI:610494
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 413)
  AUTHORS    Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
             ,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
             ,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W., Clayton,R.A.,
             ,L.M., FitzHugh,W.M., Fritchman,J.L., Geoghagen,N.S.M., Glodek,A.,
             Gnehm,C.L., Hanna,M.C., Hedblom,E., HinkleJr,P.S., Kelley,J.M.,
             Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.N., Merrick,J.M.,
             Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pellegrino,S.M.,
             Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
             Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
             Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.-J.,
             Dimke,D., Feng,P., Ferrie,A., Fischer,C., Hastings,G.A., He,W.-W.,
             Hu,J.-S., Greene,J.M., Gruber,J., Hudson,P., Kim,A., Kozak,D.L.,
             Kunsch,C., Ji,H., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei
             ,Y.-F., Wing,J., Xu,C., Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon
             ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
             Venter,J.C.
  TITLE      Initial Assessment of Human Gene Diversity and Expression Patterns
             Based Upon 83 Million Basepairs of cDNA Sequence
  JOURNAL    Nature 377, 3-174 (1995)
  MEDLINE    96026280
COMMENT      Contact: Venter, JC
             The Institute for Genomic Research
             932 Clopper Rd, Gaithersburg, MD 20878
             Tel: 3018699056
             Fax: 3018699423
             Email: tdbinfo@tdb.tigr.org
             For clone availability, additional sequence and expression
             information related to this EST, please contact the TIGR Database
             (tdbinfo@tdb.tigr.org)
             Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1..413
                     /organism="Homo sapiens"
                     /db_xref="ATCC (inhost):102309"
                     /db_xref="taxon:9606"
                     /clone_lib="Human Uterus"
                     /note="Organ: uterus"
BASE COUNT    100 a    108 c    92 g    107 t    6 others
ORIGIN
```

```
Query Match        7.1%;    Score 361;   DB 14;   Length 413;
Best Local Similarity   94.5%;   Pred. No. 1e-41;
Matches    381;   Conservative   0;   Mismatches    21;   Indels    1;   Gaps    1;

Qy  4047 CTTCCCGCAACTTACACGTGGACGACCAGATGGCTGTCATTCAGTACTCCTGGATGGGGCT 4106
         |||||||||||||||||||||||||||||| || |||||||||||||||||||||||||||
Db     1 CTTCCCGCAACTTACACGTGGACGACCAGATGGCTGTGTNATTCAGTACTCCTGGATGGGGCT 60

Qy  4107 CATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGTCAACTGCTGGATGCTCTACTT 4166
         ||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    61 CATGGTGTTTGCCATGGGCTGGCGATCCTTCACCAATGNAACTCCAGGATGCTCTACTT 120

Qy  4167 CGCCCCTGATCTGGTTTTCAATGAGTACCGCATGCACAAGTCCCGGATGTACAGCCAGTG 4226
         |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CGCCCCTNATCTGGTTTTCAATGAGTACCGCATGCACAAGTCCCGGATGTACAGCCAGTG 180

Qy  4227 TGTCCGAATGAGGCACCTCTTCAAGAGTTTGGATGGCTCCAAATCACCCCCAGGAATT 4286
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TGTCCGAATGAGGCACCTCTTCAAGAGTTTGGATGGCTCCAAATCACCCCCAGGAATT 240

Qy  4287 CCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCCAGTGCATGCGATGGGCTGAAAAATCA 4346
         |||||||||||||||||||||||||||||||||||||||||||||  || ||||||||
Db   241 CCTGTGCATGAAAGCACTGCTACTCTTCAGCATTATTCCAGTGCATGGGGATGGGCTGAAAATCA 300
```

```
BASE COUNT    151 a    209 c    183 g     85 t
ORIGIN
```

```
Query Match        7.2%;    Score 368.4;   DB 10;   Length 628;
Best Local Similarity   75.8%;   Pred. No. 7.1e-43;
Matches    502;   Conservative   0;   Mismatches    91;   Indels    69;   Gaps    1;

Qy  1804 GAGCAAGAAGAGAAGGGGAGGCGGGTAAGGGAAGTAGGTGGAAGATTCAGCCAAGCTCAAGG 1863
         ||||||| || ||||||| |  ||||| | |||||||||||||||||||||||||||||||
Db     3 GAGCAAGAAGAGGAGGAGGCCAGGATAAGGGAATTCGTGGAAGCTACAGACAAGCTCAAGG 62

Qy  1864 ATGGAAGTGCAGTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCGGTTCCAAGACCTACCGA 1923
         |||| ||||||||||||||||||||||||||||||||| | ||||| ||||||||||||
Db    63 ATGGAGGTGCAGTTAGGGCTGGGAAGGGTCTACCCTCGGCCGCCCATCCAAGACCTATCGA 122

Qy  1924 GGAGCTTTCCAGAATCTGTTCCAGAGCGTGCGCGAAGTGATCCAGAACCCGGCCCCAGG 1983
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123 GGAGCGTTCCAGAATCTGTTCCAGAGCGTGCGCGAAGCGATCCAGAACCCGGCCCCAGG 182

Qy  1984 CACCCAGGAGGCCGCGAGCCGCAGCACCTCCCGGCGCGCCAGTTGCTGCTGCTGCAGCAGCAG 2043
         ||||||| | ||||  ||| |  |  ||||||||||||||| || ||||||||||||||
Db   183 CACCCTGAGGGCCGCTAACATAGCACCTCCCGGCGCGCCTGTTT-----------------223

Qy  2044 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2103
                                              ||||||
Db   224 ------------------------------------------ACAGCAGAGG 233

Qy  2104 CAGCAAGAGACTAGCCCAGGCAGCAGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCAA 2163
         ||| |  || |  |||||| ||| || ||  ||||||||||||||||||||||||||||||
Db   234 CAGGAGACTAGCCCCCCGGCGGCGGCGGCGGCGCAGCACACTGAGGATGGTTCTCCTCCAA 293

Qy  2164 GCCCATCGTAGAGGCCCCACAGGCTACCTGCTGCTGTCCTGCGATGAGGAACAGCAACCTTCACAG 2223
         |||||||||||||||||||  |  ||||||||||||| ||||||||||||||||| |||||||
Db   294 GCCCACATCAGAGGCCCCACAGGCTACCTGGCCCTGGCCCTGGAGGAGGAACAGCAGCCTTCACAG 353

Qy  2224 CCGCAGTTCGGCCCTGGAGTGCCACCCCGAGAGAGGTTGCGCTCCCAGAGCCTGGAGCCGCCC 2283
         |||||| ||| |||||||| |  |||||| |||| ||||||||||||||||||||||||||||
Db   354 CAGCAGCAGCCTCCCGAGGGCCACCCTGAGAGCAGCTGCCTCCCCGAGCCCTGGGGCGGGCC 413

Qy  2284 GTGGGCCGCCAGCAAGGGGCTGCCGCAACGGGGGCTCGCCAGCCTCCGGACGAGGATGACTCA 2343
         || |  ||||||||  ||| |||||||  |||| ||||||| |||| ||||||||||||||
Db   414 ACCGGCTCTCCTGCCAAGGGCTGCCCGGGCAGCGGCCCCACCGCTCCTCCAGATCAGGATGACTCA 473

Qy  2344 GCTGCCCCATCCACGTTGTCCCTGTGGGCCCCCACTTTTCCCCGGCTTAAGCAGCTGCTCC 2403
         ||||||||||||||||| |||||| ||||||||||  |||||||||||||||||||||||
Db   474 GCTGCCCCATCCACGTTGTCCCTGCTGGGCCCCCACTTTCCCCAGGCTTAAGCAGCTGCTCC 533

Qy  2404 GCTGACCTTAAAGCACATCCTGAGCGAGGCCAGCAACATGCAACTCCTTCAGCAACAGCAG 2463
         ||||| ||||| ||||||||| || ||||||| ||||||  ||||||||||| |||||||||
Db   534 GCCGACACTTTGAACGAGGCCGGCACCATGCAACTCAACTTTTTAGCAGCAGCAA 593

Qy  2464 CA 2465
         ||
Db   594 CA 595
```

```
RESULT 9
T28396
LOCUS       T28396     413 bp    mRNA    linear    EST 06-SEP-1995
DEFINITION  EST41738 Human Uterus Homo sapiens cDNA 5' end similar to androgen
            receptor (Ht:1334), mRNA sequence.
```

Left column (sequence library notes):

```
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGGAAGAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 3.0 and subtraction to Rot = 100.0. Second strand
cDNA was prepared with the primer adapter of sequence [5'
GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCGCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."
```

```
Qy  4347  AAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGAACTCGATCGTATCATTGCATG  4406
          ||||||||||||| ||||| |||| || |||||| |||| ||||| ||||||||||||||
Db   301  AAAATTCTTTGATGNACTTCGGATGNACTACA-CAAGGACTCGGTCGTATCATTGCATG   359

Qy  4407  CAAAAGAAAAAATCCCACATCCTGCTCAAGACGCGTTCTACCAG  4449
          ||||||||||||||||||||||| |||| || |||||
Db   360  CAAAAGAAAAAATCCCACATCCTGNTTCAGACGGTTTTTACCAG   402
```

RESULT 10
AW619116
LOCUS   AW619116   573 bp   mRNA   linear   EST 24-MAR-2000
DEFINITION   436 MARC PBE Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   AW619116
VERSION   AW619116.1 GI:7325300
KEYWORDS   EST.
SOURCE   pig.
ORGANISM   Sus scrofa
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1 (bases 1 to 573)
AUTHORS   Smith,T.P.L., Fahrenkrug,S.C., Rohrer,G.A., Simmen,F.A., Rexroad
             ,C.E. and Keele,J.W.
TITLE   Mapping of expressed sequence tags from a porcine early embryonic
             cDNA library
JOURNAL   Anim. Genet. 32 (2), 66-72 (2001)
MEDLINE   21314990
COMMENT   Contact: Smith TPL
             USDA, ARS, US Meat Animal Research Center
             PO Box 166, Clay Center, NE 68933-0166, USA
             Tel: 402 762 4366
             Fax: 402 762 4390
             Email: smith@email.marc.usda.gov
             Single pass sequencing. Bases called and alt_trimmed with phred
             v0.980904.e. Vector identified by cross_match with the -minscore 18
             and -minmatch 12 options.
             PCR Primers
             FORWARD: GGAAACAGCTATGACCATG
             BACKWARD: GTAAAACGACGGCCAGT
             Seq primer: AATTAACCCTCACTAAAGGG.
FEATURES   Location/Qualifiers
     source 1..573
             /organism="Sus scrofa"
             /db_xref="taxon:9823"
             /clone_lib="MARC PBE"
             /tissue_type="Day 12 whole embryos"
             /lab_host="XLOLR"
             /note="Vector: pBLUESCRIPT SK-; Site_1: EcoRI; Site_2:
             XhoI; Library made from pool of embryos in spherical and
             filamentous stages of development (7.5% and 92.5%,
             respectively, of each stage) as described in Choi et al,
             Endocrinology 137, 1457-67, 1996."
BASE COUNT   125 a   199 c   162 g   87 t
ORIGIN

Query Match   7.1%;   Score 358.4;   DB 10;   Length 573;
Best Local Similarity 77.7%;   Pred. No. 2e-41;
Matches 471;   Conservative 0;   Mismatches 81;   Indels 54;   Gaps 1;

```
Qy  1863  GATGGAAGTGCAGTTAGGGCTGGGAAGGGTCTACCCTCGGCGGCGCGTCCAAGACCTACCG  1922
          | ||||||||||||||||||||||| |||||||||||||  ||||||||||||||||| ||
Db    10  GTTGGAAGTGCAGTTAGGGCTGGGGAGGGTCTACCCTCCGGCGGTCTTCCGCGCCTTTCCG   69

Qy  1923  AGGAGCTTTCCAGAATCTGTTCCAGAGCGTGCGCGCCAAGTGATCCAGAACACCGGGCCCCAG  1982
          ||| | ||||| ||| |||||||| |||| || | |||| |||||
Db    70  AGTGCTTTCCAGAACCTGTTCCAAAGTGTTACGCGAAGTGATCCGAAGAACCCAGTCCCCG   129

Qy  1983  GCACCCAGAGGGCCGCCGAGCGACCCACCTCCCGGCGGCACCTGCTCTGCTGCTGCAGCAGCA  2042
          ||| |||||||||||||||||| |||||||||||||  |||||  ||| ||||||||||||
Db   130  ACACCCTGAGGCCGGAGCGGAGCACCTCACCTCCCGGCGCCCGTTTGCAGCAGCAGCTGCA   189

Qy  2043  GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA  2102
          |||||||| ||   | | |||
```

```
Db   190  GCAGCAGGAGACCAGTCCCCGGCGGCA------------------------------  216

Qy  2103  GCAGCAAGAGAGCTAGCCCCAGGCAGCAGCAGCAGCAGCAGGGTGAGGATGGTTCTCCCCA  2162
                              ||||||||||||||||||||| | |||||||||||||||||
Db   217  ---------------GCAGCAGCAGCAGCCGAGTGAGGATGGCTCTCCCCA   255

Qy  2163  AGCCCATCGTGAGAGGCCCACAGGCTACCTGGTCCTGATGAGGAACAGCAACCTTCACA  2222
          || ||| |||||||||||||||||||| || ||| ||| ||||||||||||||||||
Db   256  AGTCCGAGACGAGAGGCCCACAGGCTACCTGGCCCTGACGAGGAAACACAGCCTTCACA   315

Qy  2223  GCGCAGTCGGCCTCGGAGTCCACCCCGAGAGAGAGGTTGCGTCCCGAGAGCCTCGGAGCCGC  2282
          ||||   |||    |||||| || ||||||||||||||||| ||||| |||||| ||  |
Db   316  ACAGCAATCAGCCCCTGAGTGCCAGCCGCGAGAGCGGCTGCACCCCAGACCGGACGAGCAGC   375

Qy  2283  CGTGGCCGCCAGCAAGGGCGTGCCCGCAGCAGCTGCCAGCAGTGACGCACCTCGGACGCAGGATGACTC  2342
          | |||||||||||||||||||||  |||||||||| |||||| || |||||||||||||||||| |
Db   376  TTCGGGCCGCCAGCAAGGGCGTGCAACAGCAGCCGTCAGCAGCGGTTCGGACGAGGATGAATTC   435

Qy  2343  AGCTGCCCCATCCACGTTGTCCCTGCGTGGGCCCCCACTTTCCCCGGCTTAAGCACGTGCTC  2402
          |||||||||||||||  |||||| |||||||| || |||||||||||| ||||||| |||||
Db   436  AGCTGCCCCATCCACCGTGTCCCGTGCGTGGGGCCCCACTTTCCCCGGGTTTAAGCAGTGCTC   495

Qy  2403  CGCTGACCTTAAAGACATCCTGAGCGAGGCCAGCACCATGCAACTCCTTCAGCAACAGCA  2462
          | ||| |||||||||||||  ||||||||||  |||||||||||| ||||||||| ||||
Db   496  CACCGATCTTAAAGACATCTTGAGCGAGGCCGGCACCATGCAACTTCTTCAGCAGCAGCA   555

Qy  2463  GCAGGA  2468
          | || |
Db   556  ACAGCA  561
```

RESULT 11
AI785019/c
LOCUS   AI785019   637 bp   mRNA   linear   EST 02-JUL-1999
DEFINITION   uj22e09.x1 Sugano mouse kidney mkia Mus musculus cDNA clone
             IMAGE:1920712 3' similar to gb:X53779 Mouse mRNA for androgen
             receptor (MOUSE)., mRNA sequence.
ACCESSION   AI785019
VERSION   AI785019.1 GI:5332735
KEYWORDS   EST.
SOURCE   house mouse.
ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 637)
AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
             Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
             ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
             ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
             Waterston,R. and Wilson,R.
TITLE   The WashU-NCI Mouse EST Project 1999
JOURNAL   Unpublished (1999)
COMMENT   Other_ESTs: uj22e09.y1
             Contact: Marra M/WashU-NCI Mouse EST Project 1999
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: mouseest@watson.wustl.edu
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             MGI:977004
             Seq primer: custom primer used
             High quality sequence stop: 518.
FEATURES   Location/Qualifiers
     source 1..637
             /organism="Mus musculus"
             /strain="C57BL"
             /db_xref="taxon:10090"
             /clone="IMAGE:1920712"
             /clone_lib="Sugano mouse kidney mkia"
             /sex="female"
             /dev_stage="adult"

```
                    /lab_host="DH10B"
                    /note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraIII
                    (CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                    was primed with an oligo(dT) primer
                    [ATGTGGCCTTTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                    ligated to a DraIII adaptor [TGTTGGCCTACTGG], digested
                    and cloned into distinct DraIII sites of the pME18S-FL3
                    vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should
                    be used to isolate the cDNA insert. Size selection was
                    performed to exclude fragments <1.5kb. Library
                    constructed by Dr. Sumio Sugano (University of Tokyo
                    Institute of Medical Science). Custom primers for
                    sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                    primer CGACCTGCAGCTCGAGCACA."
BASE COUNT      211 a    168 g    139 t     1 others
ORIGIN

Query Match        7.0%;      Score 357.8;   DB 9;    Length 637;
Best Local Similarity   81.4%;   Pred. No. 2.2e-41;
Matches    529;    Conservative    0;   Mismatches   98;   Indels   23;   Gaps    9;

Qy   4420   CCCACATCCTGCTCAGACGCGTTCTACCAGCTCACCAAGCTCCTGGACTCCGTGCAGCCT 4479
            |  |||||||||||||||||      |  ||||||||||| ||| || ||||||||
Db    634   CNCCCATCCTGCTCAGGCGGCTTCACCAGCTCACCAAG--TCTGATTTCGTGCAGCCT 577

Qy   4480   ANTGCGGAGAGCTGCATCAGTTCACTTTTGACCTGCTAATCAAGTCACACATGGTGAGC 4539
            |||| ||||||||||||||||||||  || ||||||||||||||||||| || ||||||
Db    576   ATTGCAAGAGAGCTGCATCAGTTCACTTTTTACCTGCTAATCAAGTCCCATATGGTGAGC 517

Qy   4540   GTGGACTTTCCGGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCT 4599
            |||||||||  |  ||||||||||||||||||||||||||||||||||||||||||||||
Db    516   GTGGACTTTTTCGAAATGATGGCAGAGATCATCTCTG-GCAAGTGCCCAAGATCCCTTCT 458

Qy   4600   GGGAAAGTCAAGCCCATCTATTTCCACACCCAGTGAAGCATTGGAAA-CCCTATTTCCCC 4658
            |||||||||||  | ||||||||||||||||||||||||||||     |||||  |||
Db    457   -GGAAAGTCAAGCCCATCTATTTCCACACACACAGTGAAGATTGGAAACCCCTAATACCCA 399

Qy   4659   ACCCCAGCTCATGCCCCCTTTCAGATGTCTTCTGCCTGT--TATAACTCTGCACTACTCC 4716
            |  |  |||||||||||||  ||||||||||||||||||  |||||||||||||||||||
Db    398   AAACCCACCTTGTTCCTTTCCAGATGTCTTCTGCCTGTTATATAACTCTGCACTACTTC 339

Qy   4717   TCTGCAGTGGCTT-GGGGAATTTCCTCTATTGATGTACAGTCTGTCATGGAACATGTTCCT 4775
            |||||||||||| |||||||||  ||||||||||||||||||||||||||||||| ||||
Db    338   TCTGCAGTGCCTTGGGGGAAATTTCCTCTACTGATGTACAGTCTGTCGTGAACAGGTTCCT 279

Qy   4776   GAATTCTATTTGCTGGGGCTTTTTTTTTTTTCTCGTTTTCTCTTTTTCTTTCTTCCCTCC 4835
            |  |||||||||||||||||||  ||       ||||||||||||||||||||||||||||
Db    278   CAGTTCTATTTCCTGGGCTTCTCCTT-----------CTTTTTTTTTTCTTTCTTCCCTCC 231

Qy   4836   CTATCTAACCCTCCCATGGGACCTTGCTTCAGACTTTGCTTCCCATTCTGGCGTCCTATCTGTGT 4895
            |  |  |||||||||||||||||||||||||  |||||||||||  ||||| ||||||||||||
Db    230   CTCTTTCACCCTCCCATGGGACCATTTTGAATCTGCTGCGTATTGTGGCTTCCTGCCTTTGT 171

Qy   4896   TTTGAATGGTGTGTTGTATGCCTTTAAAATCTGTGATGATCCTCATATGGCCCAGTGTCAAGT 4955
            ||||||||  ||||||||||||||||||| | |||||||||||||||  |||||||||||||
Db    170   TTTGATTTCTTGTTGTATTTCTTTGTGATTCTGTGATGATCCTCTGTGGCCCCAGTGTCAA-T 112

Qy   4956   TGTGCTTGTTTTACAGCAC--TACTCTGTGCCAGCCACACAAACGTTTACTTATCTTATGC 5013
            |||||||||||||||||||   ||||||||||||||||||||| ||||||||||| ||||||
Db    111   TGTGCTTGTTTTATAGCACTGTGCTGTGTGTCGTGGCCAACCAAACAAATGTTTACTCACCTTATGC 52

Qy   5014   CACGGGAAGTTTAGAGAGCTAAGATTATCTGGGGAAATCAAAACAAAAAA 5063
            ||||  || || |||||  | |||||||||| ||||| ||  |  ||| |
Db    51    CATGGCAAATTTAGAGAGCTATAAGTATCTGGAGAAGAAGAAACAAACAGAGA 2

RESULT 12
AA230070
LOCUS        AA230070      349 bp     mRNA      linear    EST 21-AUG-1997
DEFINITION   nc49b06.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1011443
             similar to gb:M23263 ANDROGEN RECEPTOR (HUMAN);, mRNA sequence.
ACCESSION    AA230070
VERSION      AA230070.1    GI:1852363
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 349)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.
             , Michael Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             DNA Sequencing by: Genome Systems Inc., Greg Lennon, Ph.D.
             Clone distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             www-bio.llnl.gov/bbrp/image/image.html
             Seq primer: -41m13 fwd. ET from Amersham
             High quality sequence stop: 230.
FEATURES             Location/Qualifiers
     source          1..349
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1011443"
                     /clone_lib="NCI_CGAP_Pr3"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated, total cellular RNA obtained from 5,000-10
                     ,000 microdissected cells histologically-determined to be
                     fully malignant prostate cancer cells. Double-stranded
                     cDNA was ligated to EcoRI adaptors, 5 cycles of PCR
                     applied to the cDNA with an adaptor-specific primer, and
                     the resulting PCR product subcloned into pAMP10 by the
                     UDG-cloning method (Life Technologies). Average insert
                     size is 600 bp. NOTE: Not directionally cloned. This
                     library was constructed by David Krizman."
BASE COUNT     64 a     105 c     57 g    123 t
ORIGIN

Query Match        6.8%;      Score 347.4;   DB 9;    Length 349;
Best Local Similarity   99.7%;   Pred. No. 9.2e-40;
Matches    348;    Conservative    0;   Mismatches    1;   Indels    0;   Gaps    0;

Qy   4540   GTGGACTTTCCGGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCT 4599
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1   GTGGACTTTCCGGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAAGATCCTTTCT 60

Qy   4600   GGGAAAGTCAAGCCCATCTATTTCCACACCCAGTGAAGCATTGGAAACCCTATTTCCCCA 4659
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61   GGGAAAGTCAAGCCCATCTATTTCCACACCCAGTGAAGCATTGGAAACCCTATTTCCCCA 120

Qy   4660   CCCCAGCTCATGCCCCCTTTCAGATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCT 4719
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121   CCCCAGCTCATGCCCCCTTTCAGATGTCTTCTGCCTGTTATAACTCTGCACTACTCCTCT 180

Qy   4720   GCAGTGCCTTGGGGAATTTCCTCTATTGATGTACAGTCTGTCATGAACATGTTCCTGAAT 4779
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181   GCAGTGCCTTGGGGAATTTCCTCTATTGATGTACAGTCTGTCATGAACATGTTCCTGAAT 240

Qy   4780   TCTATTTGCTGGGCTTTTTTTTCTCTTTTTCTCTTTTTTCTTTTTCTTTCTTCCCTAT 4839
            |||||||||||||||||||||||||||||||||;|||||||||||||||||||||||||
Db    241   TCTATTTGCTGGGCTTTTTTTTCTCTTTTTCTCTTTTTTCTTTTTCTTTCTTCCCTAT 300

Qy   4840   CTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCTCCTA 4888
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301   CTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCTCCTA 349
```

```
RESULT 13
AA659567/c
LOCUS       AA659567              379 bp    mRNA    linear   EST 05-NOV-1997
DEFINITION  nt63a03.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1203148
            similar to gb:M23263 ANDROGEN RECEPTOR (HUMAN);, mRNA sequence.
ACCESSION   AA659567
VERSION     AA659567.1  GI:2595721
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 379)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.
            , Michael Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 360.
FEATURES             Location/Qualifiers
     source          1..379
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1203148"
                     /clone_lib="NCI_CGAP_Pr3"
                     /sex="Male"
                     /dev_stage="45 years old"
                     /lab_host="DH10B"
                     /note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st
                     strand cDNA was primed with oligo(dT)17 on 50 ng of
                     DNAse-treated, total cellular RNA obtained from 5,000-10
                     ,000 microdissected cells histologically-determined to be
                     fully malignant prostate cancer cells. Double-stranded
                     cDNA was ligated to EcoRI adaptors, 5 cycles of PCR
                     applied to the cDNA with an adaptor-specific primer, and
                     the resulting PCR product subcloned into pAMP10 by the
                     UDG-cloning method (Life Technologies). Average insert
                     size is 600 bp. NOTE: Not directionally cloned. This
                     library was constructed by David Krizman."
BASE COUNT     130 a     67 c     95 g     87 t
ORIGIN

Query Match             6.8%;  Score 346;  DB 9;  Length 379;
Best Local Similarity   98.1%;  Pred. No. 1.4e-39;
Matches  371;  Conservative   0;  Mismatches   5;  Indels   2;  Gaps   2;

Qy   4705 CTGCACTACTCCCTGCAGTGCCTTGGGGAATTCCTCATTGCATTGATGTACAGTCTGTCATG 4764
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    376 CTGCACTACTCCCTGCAGTGCCTTGGGGAATTCCTCCATTGGTATACAGTCTGTCATG 317

Qy   4765 AACATGTTCCTGAATTCTATTGCTGGGCTTTTTTTTTTCTCTTTCTCTCCTTTCTTTTTC 4824
          |||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||
Db    316 AACATGTTCCTGAATTCTATTGCTGGGC-TTTTTTATCTCTTCCTTTCTCTTCTTTTTC 258

Qy   4825 TTCTTCCCTCCCATATCTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCT 4884
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    257 CTCTTCCCTCCCATATCTAACCCTCCCATGGCACCTTCAGACTTTGCTTCCCATTGTGGCT 198

Qy   4885 CCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAAATCTGTGATGATCCTCATATGGCC 4944
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    197 CCTATCTGTGTTTTGAATGGTGTTGTATGCCTTTAAAATCTGTGATGATCCTCATATGGCC 138
```

```
Qy   4945 CAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCAGCCACACAAAACGTTTACTT 5004
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    137 CAGTGTCAAGTTGTGCTTGTTTACAGCACTACTCTGTGCCAGCCACCACAAAACGTTTACTT 78

Qy   5005 ATCTTATGCCACGGAAGTTTAGAGAGCTAAGATTATCTGGGAAATCAAAACAAAAAAAAAC 5064
          ||||||||||||||||||||||||||||||||||||||||||||||||| || |  | | |
Db     77 ATCTTATGCCACGGAAGTTTAGAGAGCTAAGATTATCTGGGAAATCAAAAC-ACAGAC 19

Qy   5065 AAGCAAACAAAAAAAAAA 5082
          |||||||||||||||
Db     18 AAGCAAACAAAAAAAAAA 1
```

```
RESULT 14
AI315085/c
LOCUS       AI315085              536 bp    mRNA    linear   EST 17-DEC-1998
DEFINITION  uj23g04.x1 Sugano mouse kidney mkia Mus musculus cDNA clone
            IMAGE:1920822 3' similar to gb:X3779 Mouse mRNA for androgen
            receptor (MOUSE);, mRNA sequence.
ACCESSION   AI315085
VERSION     AI315085.1  GI:4030352
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 536)
AUTHORS     Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE       The WashU-HHMI Mouse EST Project
JOURNAL     Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO.63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:97114
            Seq primer: custom primer used
            High quality sequence stop: 487.
FEATURES             Location/Qualifiers
     source          1..536
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1920822"
                     /clone_lib="Sugano mouse kidney mkia"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pME18S-FL3; Site_1: DraII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a DraIII adaptor [TGTTGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert. Size selection was
                     performed to exclude fragments <1.5kb. Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science). Custom primers for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
BASE COUNT     181 a    101 c    129 g    124 t    1 others
ORIGIN

Query Match             6.3%;  Score 319;  DB 9;  Length 536;
```

```
Best Local Similarity 81.7%;  Pred. No. 7.6e-36;
Matches  446;  Conservative  0;  Mismatches  81;  Indels  19;  Gaps  6;

Qy  4531  ATGGTGAGCGTGGACTTTCCGGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAAG  4590
              |||||||||||||||||| ||||||||||| |||||||||||||||||||||||||
Db   536  ATGGTGAGCGTGGACTTTCCTGAAATGATGGCAGAGATCATCTCTGTGCAAGTGCCCAAG  477

Qy  4591  ATCCTTTCTGGAAAGTCAAGCCCATCTATTTCCACACCCAGTGAAGCATTGGAAACCCT  4650
              ||||||||||||||||||||||||||||| |||||||  |||||||||
Db   476  ATCCTTTCTGGAAAGTCAAGCCCATCTATTTCCACACCAGTGAAGATTTGGAAACTAA  417

Qy  4651  ATTTCCCCACCCCAGCTCATGCCCCCTTTCAGATGTCTTCGCCTGT--TATAACTCTGC  4708
              || |  |  || | || |||| |||||| |||||||||||||||
Db   416  ATACCCANAACCCACCTTGT-TCCCTTTCCAGATGTCTTCTGCCTGTTATATAACTCTGC  358

Qy  4709  ACTACTCCTCTGCAGTGCCTT-GGGGAATTTCCTCTATTGATGTACAGTCTGTCATGAAC  4767
              ||||||  ||||||||| ||||| |||||||| ||||||| |||||||||||  ||||||
Db   357  ACTACTTCTCTGCAGTGCCTTGGGGGAAATTCCTCNTCACTGATGTACAGTCTGTCGTGAAC  298

Qy  4768  ATGTTCCTGAATTCTATTTGCTGGGCTTTTTTCTCTTTCTCCTTTCTTTTTCTTTC  4827
              | ||||| | |||||| | |||||| ||||      |||  |||||||||||||||
Db   297  AGTTCCTCAGTTCTATTTCCTGGGCTTCTCCTT--------CTTTTTTTTCTTC  250

Qy  4828  TTCCCTCCCTATCTAACCCTCCCATGGCACCTTCAGACTTGCTTCCCATTGTGGCTCCT  4887
              |||||| ||||||| | |  |||||||||| ||| ||||| |||||||||||||||
Db   249  TTCCCTCCCTCTTTCACCCTCCCATGGCACACCCATTTGAAATCTGCTGCGTATTGTGGCTCCT  190

Qy  4888  ATCTGTGTTTTGAATGGTGTGTTGTATGCCTTTAAATCTGTGATGATCCTCATATGGCCCAG  4947
              || ||||||  | ||||||||| ||| ||| || |||||||||  | |||||| |||||||
Db   189  GCCTTTGTTTGTTGATTTCTGTGTATTTCTTTGAATCTGTGATGATCCCTCTTGTGGCCCAG  130

Qy  4948  TGTCAAGTTGTGTCGTTGTTTACAGCAC--TACTCTGTGCCAGCCACACAAAACGTTTACTTA  5005
              | |||| ||||||||||||||| ||| ||  |||||||| ||||||||||||||||||||
Db   129  TGTCAA--TTGTGCTTGTTTATAGACACTGTGTGTGTGCCAACCAAGCAAATGTTTACTCA  71

Qy  5006  TCTTATGCCACGGAAGTTTAGAGAGCTAAGATTATCTGGGAAATCAAAACAAAAAACA  5065
              ||||||||||| | |||| | ||||| || | |||| |   |   || |
Db    70  CCTTATGCCATGGCAAATTTAAAAAGCTATAAGTATCTGGAGAGAAAAACAAACAGAGA  11

Qy  5066  AGCAAA  5071
              | |||
Db    10  ATAAAA  5
```

RESULT 15
AZ017963
LOCUS   AZ017963   498 bp   DNA   linear   GSS 25-FEB-2000
DEFINITION RPCI-23-259K23.TV RPCI-23 Mus musculus genomic clone RPCI-23-259K23
            , DNA sequence.
ACCESSION   AZ017963
VERSION   AZ017963.1 GI:7093347
KEYWORDS   GSS.
SOURCE   house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 498)
AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
TITLE   Mouse BAC End Sequences from Library RPCI-23
JOURNAL   Unpublished (1999)
COMMENT   Other_GSSs: RPCI-23-259K23.TJ
            Contact: Shaying Zhao
            Email: szhao@tigr.org
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Research Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 259 row: K column: 23
Seq primer: T7
Class: BAC ends.

```
FEATURES       Location/Qualifiers
     source    1. .498
               /organism="Mus musculus"
               /strain="C57BL/6J"
               /db_xref="taxon:10090"
               /clone="RPCI-23-259K23"
               /clone_lib="RPCI-23"
               /sex="Female"
               /lab_host="DH10B"
               /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
               EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
               brain genomic DNA was isolated and partially digested
               with a combination of EcoRI and EcoRI Methylase.  Size
               selected DNA was cloned into the pBACe3.6 vector at the
               EcoRI sites.  The ligation products were transformed into
               DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT      102 a   173 c   147 g    75 t    1 others
ORIGIN
```

```
Query Match           5.8%;    Score 293.2;  DB 17;  Length 498;
Best Local Similarity 73.9%;   Pred. No. 3.5e-32;
Matches 419;  Conservative  0;  Mismatches  79;  Indels  69;  Gaps  1;

Qy  1865  TGGAAGTGCAGTTAGGGCTGGGAAGGGTCTACCCTCGGCGCCGCCGGTCCAAGACCTACCGAG  1924
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db     1  TGGAGGTGCAGTTAGGGCTGGGAAGGGTCTCACCCGCCGGTCATCCAAGACCTATCGAG  60

Qy  1925  GAGCTTTCCAGAATCTGTTCCAGACGCTGCGCGAAGTGACGCCGGGCCCCAGGC  1984
              |||| ||||||||||||||||||||||||||||||| |||||||||||||||||
Db    61  GAGCGTTCCAGAATCTGTTCCAGACGCTGCGCGAAGCGATCCAGAACCCGGGCCCCAGGC  120

Qy  1985  ACCCAGAGGGCCGGAGCGAGCAGCACCTCCCGGCGCCAGTTGCTGCTGCGCAGCAGCAGC  2044
              ||||||||||| | |||||||||||||||||||||||||||||| |||
Db   121  ACCCTGAGGGCCGCTAACATAGCACCTCCCGGGCGCCGGCTGTTT------------  160

Qy  2045  AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC  2104
                                                       |||||| |||||
Db   161  ------------------------------------------ACAGCAGAGGC  171

Qy  2105  AGCAAGAGACTAGCCCAGGCAGCAGCAGCAGCAGGTGAGGATGGTTCTCCCAAG  2164
              |||  |||  || ||||||| || | |||||||||| |||||||||||||
Db   172  AGGAGTCNAGCCCCCGGCGGCGGTTGCGGCAGCAGCACACTGAGGATGGTTCTCCTCAAG  231

Qy  2165  CCCATCGTAGAGGCCCCAGAGGTCTACCTGGTCTCTGGATGAGGAACAGCAGCACCTTCACAGC  2224
              |||| ||||||||||||||||||||||||||||| ||||||||||||||
Db   232  CCCACATCAGAGGCCCCACAGGTCTACCTGGCCTCGGCGGAGGAACAGCAGCCTTCACAGC  291

Qy  2225  CGCAGTCGGCCCTGGAGTGCACCCCGGAGCAGAGGTTGCTGCAGCAGCCTGGAGCCGCGCG  2284
              |||  || |||  |||||||||||||||||||||||||||| || |||||||||
Db   292  AGCAGCCGCAGCCTCCGAGGGCCACCCTGAGAGCAGCAGCAGCTGCTCCCCGAGCTCGGGGCGGCCA  351

Qy  2285  TGGCCGCCAGCAAGGGGTCTGCCGCCAGCAGTCGCCAGCAGCTGCCGGCAGGAGGATGACTCAG  2344
              || |||||| ||| || ||||| |||||||||  ||| ||||||||||||||||||||
Db   352  CCGCTCCTTGCCAGGGGGTGCGCCAGCAGCCACCAGCTCCTCCAGATCAGGATGACTCAG  411

Qy  2345  CTGCCCCATCCACGTTGCTCCCTGCTGGGCCCCACTTTCCCCGGCTTAAGCAGCTGCTCCG  2404
              |||||||||||||||||||  |||||||| ||||| ||||||  |||| |||| |||||
Db   412  CTGCCCCATCCACGTGTCCCTGCTGGGCCCCACTTTCCCAGGCTTAAGCAGCTGCTCCG  471

Qy  2405  CTGACCTTAAAGACATCCTGAGCGAGG  2431
              | | |||||| |||| |||| ||||
Db   472  CCGACATTAAAGACATTTGAACGAGG  498
```

Search completed: May 8, 2003, 20:23:23
Job time : 6158 secs